SEQUENCE PROTOCOL

(1) GENERAL DETAILS :

   (i)APPLICANTS
       (A) NAME: Forschungszentrum Juelich GmbH
       (B) STREET.: Postfach 1913
       (C)LOCALE:Juelich
       (E)COUNTRY: GERMANY
       (F) ZIP CODE   .: 52425:

   (ii)DESIGNATION OF THE INVENTION:Pyruvate Carboxylase

   (iii)NUMBER OF   SEQUENCES: 2

   (iv) COMPUTER-READABLE FORM :
       (A) DATA CATEGORY:Floppy disk
       (B) COMPUTER: IBM PC compatible
       (C)OPERATING SYSTEM: PC-DOS/MS-DOS
       (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) DETAILS TO SEQ ID NO: 1:

   (i) SEQUENCE CHARACTERISTICS:
       (A)LENGTH : 3728 Base PAIRS
       (B)TYPE : Nucleotide
       (C)STRAND SHAPE : Single strand
       (D) TOPOLOGY : linear

   (ii)TYPE OF MOLECULES : Genom DNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CGCAACCGTG CTTGAAGTCG TGCAGGTCAG GGGAGTGTTG CCCGAAAACA TTGAGAGGAA   60

AACAAAAACC GATGTTTGAT TGGGGGAATC GGGGGTTACG ATACTAGGAC GCAGTGACTG   120

CTATCACCCT TGGCGGTCTC TTGTTGAAAG GAATAATTAC TCTAGTGTCG ACTCACACAT   180

CTTCAACGCT TCCAGCATTC AAAAAGATCT TGGTAGCAAA CCGCGGCGAA ATCGCGGTCC   240

GTGCTTTCCG TGCAGCACTC GAAACCGGTG CAGCCACGGT AGCTATTTAC CCCCGTGAAG   300

ATCGGGGATC ATTCCACCGC TCTTTTGCTT CTGAAGCTGT CCGCATTGGT ACCGAAGGCT   360

CACCAGTCAA GGCGTACCTG GACATCGATG AAATTATCGG TGCAGCTAAA AAAGTTAAAG   420

CAGATGCCAT TTACCCGGGA TACGGCTTCC TGTCTGAAAA TGCCCAGCTT GCCCGCGAGT   480

2

21437

```
GTGCGGAAAA CGGCATTACT TTTATTGGCC CAACCCCAGA GGTTCTTGAT CTCACCGGTG      540

ATAAGTCTCG CGCGGTAACC GCCGCGAAGA AGGCTGGTCT GCCAGTTTTG GCGGAATCCA      600

CCCCGAGCAA AAACATCGAT GAGATCGTTA AAAGCGCTGA AGGCCAGACT TACCCCATCT      660

TTGTGAAGGC AGTTGCCGGT GGTGGCGGAC GCGGTATGCG TTTTGTTGCT TCACCTGATG      720

AGCTTCGCAA ATTAGCAACA GAAGCATCTC GTGAAGCTGA AGCGGCTTTC GGCGATGGCG      780

CGGTATATGT CGAACGTGCT GTGATTAACC CTCAGCATAT TGAAGTGCAG ATCCTTGGCG      840

ATCACACTGG AGAAGTTGTA CACCTTTATG AACGTGACTG CTCACTGCAG CGTCGTCACC      900

AAAAAGTTGT CGAAATTGCG CCAGCACAGC ATTTGGATCC AGAACTGCGT GATCGCATTT      960

GTGCGGATGC AGTAAAGTTC TGCCGCTCCA TTGGTTACCA GGGCGCGGGA ACCGTGGAAT     1020

TCTTGGTCGA TGAAAAGGGC AACCACGTCT TCATCGAAAT GAACCCACGT ATCCAGGTTG     1080

AGCACACCGT GACTGAAGAA GTCACCGAGG TGGACCTGGT GAAGGCGCAG ATGCGCTTGG     1140

CTGCTGGTGC AACCTTGAAG GAATTGGGTC TGACCCAAGA TAAGATCAAG ACCCACGGTG     1200

CAGCACTGCA GTGCCGCATC ACCACGGAAG ATCCAAACAA CGGCTTCCGC CCAGATACCG     1260

GAACTATCAC CGCGTACCGC TCACCAGGCG GAGCTGGCGT TCGTCTTGAC GGTGCAGCTC     1320

AGCTCGGTGG CGAAATCACC GCACACTTTG ACTCCATGCT GGTGAAAATG ACCTGCCGTG     1380

GTTCCGACTT TGAAACTGCT GTTGCTCGTG CACAGCGCGC GTTGGCTGAG TTCACCGTGT     1440

CTGGTGTTGC AACCAACATT GGTTTCTTGC GTGCGTTGCT GCGGGAAGAG GACTTCACTT     1500

CCAAGCGCAT CGCCACCGGA TTCATTGCCG ATCACCCGCA CCTCCTTCAG GCTCCACCTG     1560

CTGATGATGA GCAGGGACGC ATCCTGGATT ACTTGGCAGA TGTCACCGTG AACAAGCCTC     1620

ATGGTGTGCG TCCAAAGGAT GTTGCAGCTC CTATCGATAA GCTGCCTAAC ATCAAGGATC     1680

TGCCACTGCC ACGCGGTTCC CGTGACCGCC TGAAGCAGCT TGGCCCAGCC GCGTTTGCTC     1740

GTGATCTCCG TGAGCAGGAC GCACTGGCAG TTACTGATAC CACCTTCCGC GATGCACACC     1800

AGTCTTTGCT TGCGACCCGA GTCCGCTCAT TCGCACTGAA GCCTGCGGCA GAGGCCGTCG     1860

CAAAGCTGAC TCCTGAGCTT TTGTCCGTGG AGGCCTGGGG CGGCGCGACC TACGATGTGG     1920

CGATGCGTTT CCTCTTTGAG GATCCGTGGG ACAGGCTCGA CGAGCTGCGC GAGGCGATGC     1980

CGAATGTAAA CATTCAGATG CTGCTTCGCG CCGCAACAC CGTGGGATAC ACCCCGTACC     2040

CAGACTCCGT CTGCCGCGCG TTTGTTAAGG AAGCTGCCAG CTCCGGCGTG GACATCTTCC     2100
```

21437

GCATCTTCGA CGCGCTTAAC GACGTCTCCC AGATGCGTCC AGCAATCGAC GCAGTCCTGG   2160

AGACCAACAC CGCGGTAGCC GAGGTGGCTA TGGCTTATTC TGGTGATCTC TCTGATCCAA   2220

ATGAAAAGCT CTACACCCTG GATTACTACC TAAAGATGGC AGAGGAGATC GTCAAGTCTG   2280

GCGCTCACAT CTTGGCCATT AAGGATATGG CTGGTCTGCT TCGCCCAGCT GCGGTAACCA   2340

AGCTGGTCAC CGCACTGCGC CGTGAATTCG ATCTGCCAGT GCACGTGCAC ACCCACGACA   2400

CTGCGGGTGG CCAGCTGGCA ACCTACTTTG CTGCAGCTCA AGCTGGTGCA GATGCTGTTG   2460

ACGGTGCTTC CGCACCACTG TCTGGCACCA CCTCCCAGCC ATCCCTGTCT GCCATTGTTG   2520

CTGCATTCGC GCACACCCGT CGCGATACCG GTTTGAGCCT CGAGGCTGTT TCTGACCTCG   2580

AGCCGTACTG GGAAGCAGTG CGCGGACTGT ACCTGCCATT TGAGTCTGGA ACCCCAGGCC   2640

CAACCGGTCG CGTCTACCGC CACGAAATCC CAGGCGGACA GTTGTCCAAC CTGCGTGCAC   2700

AGGCCACCGC ACTGGGCCTT GCGGATCGTT TCGAACTCAT CGAAGACAAC TACGCAGCCG   2760

TTAATGAGAT GCTGGGACGC CCAACCAAGG TCACCCCATC CTCCAAGGTT GTTGGCGACC   2820

TCGCACTCCA CCTCGTTGGT GCGGGTGTGG ATCCAGCAGA CTTTGCTGCC GATCCACAAA   2880

AGTACGACAT CCCAGACTCT GTCATCGCGT TCCTGCGCGG CGAGCTTGGT AACCCTCCAG   2940

GTGGCTGGCC AGAGCCACTG CGCACCCGCG CACTGGAAGG CCGCTCCGAA GGCAAGGCAC   3000

CTCTGACGGA AGTTCCTGAG GAAGAGCAGG CGCACCTCGA CGCTGATGAT TCCAAGGAAC   3060

GTCGCAATAG CCTCAACCGC CTGCTGTTCC CGAAGCCAAC CGAAGAGTTC CTCGAGCACC   3120

GTCGCCGCTT CGGCAACACC TCTGCGCTGG ATGATCGTGA ATTCTTCTAC GGCCTGGTCG   3180

AAGGCCGCGA GACTTTGATC CGCCTGCCAG ATGTGCGCAC CCCACTGCTT GTTCGCCTGG   3240

ATGCGATCTC TGAGCCAGAC GATAAGGGTA TGCGCAATGT TGTGGCCAAC GTCAACGGCC   3300

AGATCCGCCC AATGCGTGTG CGTGACCGCT CCGTTGAGTC TGTCACCGCA ACCGCAGAAA   3360

AGGCAGATTC CTCCAACAAG GGCCATGTTG CTGCACCATT CGCTGGTGTT GTCACCGTGA   3420

CTGTTGCTGA AGGTGATGAG GTCAAGGCTG GAGATGCAGT CGCAATCATC GAGGCTATGA   3480

AGATGGAAGC AACAATCACT GCTTCTGTTG ACGGCAAAAT CGATCGCGTT GTGGTTCCTG   3540

CTGCAACGAA GGTGGAAGGT GGCGACTTGA TCGTCGTCGT TTCCTAAACC TTTCTGTAAA   3600

AAGCCCCGCG TCTTCCTCAT GGAGGAGGCG GGGCTTTTTG GCCAAGATG GGAGATGGGT   3660

GAGTTGGATT TGGTCTGATT CGACACTTTT AAGGGCAGAG ATTTGAAGAT GGAGACCAAG   3720

21437

GCTCAAAG   3728

(2) DETAILS TO SEQ ID NO: 2:

    (i) SEQUENCE CHARACTERISTICS:
        (A)LENGTH: 1140 Aminosäuren
        (B) TYPE: Aminosäure
        (C) STRAND SHAPE: single strand
        (D) TOPOLOGY: linear

   (ii) TYPE OF MOLECULE: Protein

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Ser Thr His Thr Ser Ser Thr Leu Pro Ala Phe Lys Lys Ile Leu
1               5                   10                  15

Val Ala Asn Arg Gly Glu Ile Ala Val Arg Ala Phe Arg Ala Ala Leu
            20                  25                  30

Glu Thr Gly Ala Ala Thr Val Ala Ile Tyr Pro Arg Glu Asp Arg Gly
        35                  40                  45

Ser Phe His Arg Ser Phe Ala Ser Glu Ala Val Arg Ile Gly Thr Glu
    50                  55                  60

Gly Ser Pro Val Lys Ala Tyr Leu Asp Ile Asp Glu Ile Ile Gly Ala
65                  70                  75                  80

Ala Lys Lys Val Lys Ala Asp Ala Ile Tyr Pro Gly Tyr Gly Phe Leu
                85                  90                  95

Ser Glu Asn Ala Gln Leu Ala Arg Glu Cys Ala Glu Asn Gly Ile Thr
                100                 105                 110

Phe Ile Gly Pro Thr Pro Glu Val Leu Asp Leu Thr Gly Asp Lys Ser
            115                 120                 125

Arg Ala Val Thr Ala Ala Lys Lys Ala Gly Leu Pro Val Leu Ala Glu
        130                 135                 140

Ser Thr Pro Ser Lys Asn Ile Asp Glu Ile Val Lys Ser Ala Glu Gly
145                 150                 155                 160

Gln Thr Tyr Pro Ile Phe Val Lys Ala Val Ala Gly Gly Gly Gly Arg
                165                 170                 175

Gly Met Arg Phe Val Ala Ser Pro Asp Glu Leu Arg Lys Leu Ala Thr
            180                 185                 190

Glu Ala Ser Arg Glu Ala Glu Ala Ala Phe Gly Asp Gly Ala Val Tyr
```

21437

```
              195                   200                   205

  Val Glu Arg Ala Val Ile Asn Pro Gln His Ile Glu Val Gln Ile Leu
     210                   215                   220

  Gly Asp His Thr Gly Glu Val Val His Leu Tyr Glu Arg Asp Cys Ser
  225                   230                   235                   240

  Leu Gln Arg Arg His Gln Lys Val Val Glu Ile Ala Pro Ala Gln His
              245                   250                   255

  Leu Asp Pro Glu Leu Arg Asp Arg Ile Cys Ala Asp Ala Val Lys Phe
              260                   265                   270

  Cys Arg Ser Ile Gly Tyr Gln Gly Ala Gly Thr Val Glu Phe Leu Val
              275                   280                   285

  Asp Glu Lys Gly Asn His Val Phe Ile Glu Met Asn Pro Arg Ile Gln
     290                   295                   300

  Val Glu His Thr Val Thr Glu Glu Val Thr Glu Val Asp Leu Val Lys
  305                   310                   315                   320

  Ala Gln Met Arg Leu Ala Ala Gly Ala Thr Leu Lys Glu Leu Gly Leu
              325                   330                   335

  Thr Gln Asp Lys Ile Lys Thr His Gly Ala Ala Leu Gln Cys Arg Ile
              340                   345                   350

  Thr Thr Glu Asp Pro Asn Asn Gly Phe Arg Pro Asp Thr Gly Thr Ile
              355                   360                   365

  Thr Ala Tyr Arg Ser Pro Gly Gly Ala Gly Val Arg Leu Asp Gly Ala
     370                   375                   380

  Ala Gln Leu Gly Gly Glu Ile Thr Ala His Phe Asp Ser Met Leu Val
  385                   390                   395                   400

  Lys Met Thr Cys Arg Gly Ser Asp Phe Glu Thr Ala Val Ala Arg Ala
              405                   410                   415

  Gln Arg Ala Leu Ala Glu Phe Thr Val Ser Gly Val Ala Thr Asn Ile
              420                   425                   430

  Gly Phe Leu Arg Ala Leu Leu Arg Glu Glu Asp Phe Thr Ser Lys Arg
              435                   440                   445

  Ile Ala Thr Gly Phe Ile Ala Asp His Pro His Leu Leu Gln Ala Pro
     450                   455                   460

  Pr  Ala Asp Asp Glu Gln Gly Arg Il  Leu Asp Tyr Leu Ala Asp Val
  465                   470                   475                   480

  Thr Val Asn Lys Pro His Gly Val Arg Pro Lys Asp Val Ala Ala Pro
```

21437

                    485   490   495

Ile Asp Lys Leu Pro Asn Ile Lys Asp Leu Pro Leu Pro Arg Gly Ser
            500   505   510

Arg Asp Arg Leu Lys Gln Leu Gly Pro Ala Ala Phe Ala Arg Asp Leu
            515   520   525

Arg Glu Gln Asp Ala Leu Ala Val Thr Asp Thr Thr Phe Arg Asp Ala
        530   535   540

His Gln Ser Leu Leu Ala Thr Arg Val Arg Ser Phe Ala Leu Lys Pro
545   550   555   560

Ala Ala Glu Ala Val Ala Lys Lys Thr Pro Glu Leu Leu Ser Val Glu
                565   570   575

Ala Trp Gly Gly Ala Thr Tyr Asp Val Ala Met Arg Phe Leu Phe Glu
                580   585   590

Asp Pro Trp Asp Arg Leu Asp Glu Leu Arg Glu Ala Met Pro Asn Val
            595   600   605

Asn Ile Gln Met Leu Leu Arg Gly Arg Asn Thr Val Gly Tyr Thr Pro
    610   615   620

Tyr Pro Asp Ser Val Cys Arg Ala Phe Val Lys Glu Ala Ala Ser Ser
625   630   635   640

Gly Val Asp Ile Phe Arg Ile Phe Asp Ala Leu Asn Asp Val Ser Gln
                645   650   655

Met Arg Pro Ala Ile Asp Ala Val Leu Glu Thr Asn Thr Ala Val Ala
            660   665   670

Glu Val Ala Met Ala Tyr Ser Gly Asp Leu Ser Asp Pro Asn Glu Lys
        675   680   685

Leu Tyr Thr Leu Asp Tyr Tyr Leu Lys Met Ala Glu Glu Ile Val Lys
    690   695   700

Ser Gly Ala His Ile Leu Ala Ile Lys Asp Met Ala Gly Leu Leu Arg
705   710   715   720

Pro Ala Ala Val Thr Lys Leu Val Thr Ala Leu Arg Arg Glu Phe Asp
                725   730   735

Leu Pro Val His Val His Thr His Asp Thr Ala Gly Gly Gln Leu Ala
            740   745   750

Thr Tyr Phe Ala Ala Ala Gln Ala Gly Ala Asp Ala Val Asp Gly Ala
        755   760   765

Ser Ala Pro Leu Ser Gly Thr Thr Ser Gln Pro Ser Leu Ser Ala Ile

21437

```
        770                     775                     780
   Val Ala Ala Phe Ala His Thr Arg Arg Asp Thr Gly Leu Ser Leu Glu
   785                     790                     795                     800

   Ala Val Ser Asp Leu Glu Pro Tyr Trp Glu Ala Val Arg Gly Leu Tyr
        -           805                     810                     815

   Leu Pro Phe Glu Ser Gly Thr Pro Gly Pro Thr Gly Arg Val Tyr Arg
                  820                     825                     830

   His Glu Ile Pro Gly Gly Gln Leu Ser Asn Leu Arg Ala Gln Ala Thr
            835                     840                     845

   Ala Leu Gly Leu Ala Asp Arg Phe Glu Leu Ile Glu Asp Asn Tyr Ala
        850                     855                     860

   Ala Val Asn Glu Met Leu Gly Arg Pro Thr Lys Val Thr Pro Ser Ser
   865                     870                     875                     880

   Lys Val Val Gly Asp Leu Ala Leu His Leu Val Gly Ala Gly Val Asp
                  885                     890                     895

   Pro Ala Asp Phe Ala Ala Asp Pro Gln Lys Tyr Asp Ile Pro Asp Ser
                  900                     905                     910

   Val Ile Ala Phe Leu Arg Gly Glu Leu Gly Asn Pro Pro Gly Gly Trp
            915                     920                     925

   Pro Glu Pro Leu Arg Thr Arg Ala Leu Glu Gly Arg Ser Glu Gly Lys
        930                     935                     940

   Ala Pro Leu Thr Glu Val Pro Glu Glu Glu Gln Ala His Leu Asp Ala
   945                     950                     955                     960

   Asp Asp Ser Lys Glu Arg Arg Asn Ser Leu Asn Arg Leu Leu Phe Pro
                  965                     970                     975

   Lys Pro Thr Glu Glu Phe Leu Glu His Arg Arg Arg Phe Gly Asn Thr
                  980                     985                     990

   Ser Ala Leu Asp Asp Arg Glu Phe Phe Tyr Gly Leu Val Glu Gly Arg
                  995                     1000                    1005

   Glu Thr Leu Ile Arg Leu Pro Asp Val Arg Thr Pro Leu Leu Val Arg
            1010                    1015                    1020

   Leu Asp Ala Ile Ser Glu Pro Asp Asp Lys Gly Met Arg Asn Val Val
   1025                    1030                    1035                    1040

   Ala Asn Val Asn Gly Gln Ile Arg Pro Met Arg Val Arg Asp Arg Ser
                  1045                    1050                    1055

   Val Glu Ser Val Thr Ala Thr Ala Glu Lys Ala Asp Ser Ser Asn Lys
```

21437

```
                    1060                    1065                    1070

      Gly His Val Ala Ala Pro Phe Ala Gly Val Val Thr Val Thr Val Ala
              1075                    1080                    1085

      Glu Gly Asp Glu Val Lys Ala Gly Asp Ala Val Ala Ile Ile Glu Ala
              1090                    1095                    1100

      Met Lys Met Glu Ala Thr Ile Thr Ala Ser Val Asp Gly Lys Ile Asp
      1105                    1110                    1115                    1120

      Arg Val Val Val Pro Ala Ala Thr Lys Val Glu Gly Gly Asp Leu Ile
                    1125                    1130                    1135

      Val Val Val Ser
                1140
```